

TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULASE
WHICH DEGRADE LIGNIN AND CELLULOSE
TO FERMENTABLE SUGARS

CROSS-REFERENCE TO RELATED APPLICATION

This application claims priority to provisional application Serial No. 60/242,408, filed October 20, 2000.

5

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

None.

10 Reference to a "Nucleotide/Amino Acid Sequence Listing Appendix submitted on a Compact Disc"

The application contains nucleotide and amino acid sequences which are identified with SEQ ID NOS. A compact disc is provided which contains the Sequence Listings for the sequences. The Sequence Listing on the compact disc and is identical to the paper copy of the Sequence Listing provided with the application.

BACKGROUND OF THE INVENTION

20 (1) Field of the Invention

The present invention relates to transgenic plants which after harvest degrade the lignin and cellulose therein to fermentable sugars which can further be fermented to ethanol or other products. In particular, the transgenic plants comprise ligninase and cellulase genes from microbes operably linked to a DNA encoding a signal peptide which targets the fusion polypeptide produced therefrom to an organelle of the plant, in particular the chloroplasts. When the transgenic plants are harvested, the plants are ground to release the ligninase and cellulase which then

degrade the lignin and cellulose of the transgenic plants to produce the fermentable sugars.

(2) Description of Related Art

5 If human economies are to become more sustainable, then it is imperative that humans learn how to use the solar energy that is carbon-fixed in plant biomass to meet a larger fraction of energy and raw material needs. About 180 billion tons of new plant matter (biomass) is produced annually worldwide. Thus, 10 about 30 tons of plant matter per person is produced every year. In North America, about three tons of plant matter is used per person every year. That means that the energy value of naturally produced biomass is 15 equivalent to ten times the total human use of all types of energy. However, because of the difficulty in extracting the energy from plant biomass, most of the energy potential of the biomass goes unused.

20 At present, the United States produces ethanol from starch produced in corn grain using amylase enzymes to degrade the starch to fermentable sugars. Much of the ethanol that is produced from corn grain is exported to Brazil where it is efficiently used to power transportation vehicles. In general, while the corn 25 grain is used in the production of ethanol, the remainder of the corn biomass, i.e., the leaves and stalks, is seldom unused because of the cost in degrading the leaves and stalks comprising lignins and cellulose, generally in the form of lignocellulose, to 30 fermentable sugars. The lignocellulose in the stalks and leaves of corn biomass represents a tremendous source of untapped energy that goes unused because of the difficulty and cost of converting it to fermentable sugars.

35 Currently, there are four technologies

available to convert cellulose to fermentable sugars. These are concentrated acid hydrolysis, dilute acid hydrolysis, biomass gasification and fermentation, and enzymatic hydrolysis.

5 Concentrated acid hydrolysis is based on concentrated acid de-crystallization of cellulose followed by dilute acid hydrolysis to sugars at near theoretical yields. Separation of acid from sugars, acid recovery, and acid re-concentration are critical
10 unit operations. The concentrated sulfuric acid process has been commercialized in the past, particularly in the former Soviet Union, Germany, and Japan. However, these processes were only successful during times of national
15 crisis, when economic competitiveness of ethanol production could be ignored.

20 Dilute acid hydrolysis occurs in two stages to maximize sugar yields from the hemicellulose and cellulose fractions of biomass. The first stage is operated under milder conditions to hydrolyze hemicellulose, while the second stage is optimized to hydrolyze the more resistant cellulose fraction. Liquid hydrolyzates are recovered from each stage, neutralized, and fermented to ethanol. As indicated earlier, Germany, Japan, and Russia have operated dilute acid
25 hydrolysis percolation plants off and on over the past 50 years. However, the technology remains non-competitive for the conversion of cellulose to fermentable sugars for production of ethanol.

30 In biomass gasification and fermentation, biomass is converted to a synthesis gas, which consists primarily of carbon monoxide, carbon dioxide, and hydrogen) via a high temperature gasification process. Anaerobic bacteria are then used to convert the synthesis gas into ethanol.

35 In early processes embracing enzymatic

hydrolysis of biomass to ethanol, the acid hydrolysis step was replaced with an enzyme hydrolysis step. This process scheme was often referred to as separate hydrolysis and fermentation (SHF) (Wilke et al., 5 Biotechnol. Bioengin. 6: 155-175 (1976)). In SHF, pretreatment of the biomass is required to make the cellulose more accessible to the enzymes. Many pretreatment options have been considered, including both thermal and chemical steps. The most important 10 process improvement made for the enzymatic hydrolysis of biomass was the introduction of simultaneous saccharification and fermentation (SSF) U.S. Patent 3,990,944 to Gauss et al. and U.S. Patent 3,990,945 to Huff et al.). This process scheme reduced the number of 15 reactors involved by eliminating the separate hydrolysis reactor and, more importantly, avoiding the problem of product inhibition associated with enzymes.

In the presence of glucose, β -glucosidase stops hydrolyzing cellobiose. The build up of cellobiose, in turn, shuts down cellulose degradation. 20 In the SSF process scheme, cellulase enzyme and fermenting microbes are combined. As sugars are produced by the enzymes, the fermentative organisms convert them to ethanol. The SSF process has, more 25 recently, been improved to include the co-fermentation of multiple sugar substrates in a process known as simultaneous saccharification and co-fermentation (SSCF) (www.ott.doe.gov/biofuels/enzymatic.html).

While cellulase enzymes are already 30 commercially available for a variety of applications. Most of these applications do not involve extensive hydrolysis of cellulose. For example, the textile industry applications for cellulases require less than 1% hydrolysis. Ethanol production, by contrast, 35 requires nearly complete hydrolysis. In addition, most

of the commercial applications for cellulase enzymes represent higher value markets than the fuel market. For these reasons, enzymatic hydrolysis of biomass to ethanol remains non-competitive.

5 However, while the above processes have focused on converting cellulose to fermentable sugars or other products, much of the cellulose in plant biomass is in the form of lignocellulose. Lignin is a complex macromolecule consisting of aromatic units with several
10 types of inter-unit linkages. In the plant, the lignin physically protects the cellulose polysaccharides in complexes called lignocellulose. To degrade the cellulose in the lignocellulose complexes, the lignin must first be degraded. While lignin can be removed in
15 chemi-mechanical processes that free the cellulose for subsequent conversion to fermentable sugars, the chemi-mechanical processes are inefficient. Ligninase and cellulase enzymes, which are produced by various microorganisms, have been used to convert the lignins and cellulose, respectively, in plant biomass to fermentable sugars. However, the cost for these enzymes is expensive, about six dollars a pound. As long as the cost to degrade plant biomass remains expensive, the
20 energy locked up in the plant biomass will largely
25 remain unused.

An attractive means for reducing the cost of degrading plant biomass is to make transgenic plants that contain cellulases. For example, WO 98/11235 to Lebel *et al.* discloses transgenic plants that express cellulases in the chloroplasts of the transgenic plants or transgenic plants wherein the cellulases are targeted to the chloroplasts. Preferably, the cellulases are operably linked to a chemically-inducible promoter to restrict expression of the cellulase to an appropriate time. However, because a substantial portion of the
30
35

cellulose in plants is in the form of lignocellulose, extracts from the transgenic plants are inefficient at degrading the cellulose in the lignocellulose.

U.S. Patent No. 5,981,835 to Austin-Phillips et al. discloses transgenic tobacco and alfalfa which express the cellulases E2, or E3 from *Thermomonospora fusca*. The genes encoding the E2 or E3, which were modified to remove their leader sequence, were placed under the control of a constitutive promoter and stably integrated into the plant genome. Because the leader sequence had been removed, the E2 or E3 product preferentially accumulated in the cytoplasm of the transgenic plants. However, because the cellulase can leak out of the cytoplasm and into the cell wall where it can degrade cellulose in the cell wall, the growth of the transgenic plants can be impaired.

U.S. Patent No. 6,013,860 to Himmel et al. discloses transgenic plants which express the cellulase E1 from *Acidothermus cellulolyticus*. The gene encoding E1, which was modified to remove the leader region, was placed under the control of a plastid specific promoter and preferably integrated into the plastid genome. Because the leader sequence had been removed, the E1 product accumulated in the plastid.

While the above transgenic plants are an improvement, accumulation of cellulytic enzymes in the cytoplasm of a plant is undesirable since there is the risk that the cellulase can leak out from the cytoplasm and injure the plant. For example, research has shown that plants such as the avocado, bean, pepper, peach, poplar, and orange also contain cellulase genes, which are activated by ethylene during ripening and leaf and fruit abscission. Therefore, transgenic plants which contain large quantities of cellulase in the cytoplasm are particularly prone to damage. Furthermore, the

cellulases accumulate in all tissues of the plant which can be undesirable. Restriction of cellulase expression to plastids is desirable because it reduces the risk of plant damage due to the cellulases leaking from the cell.

5 However, for most crop plants, it has been difficult to develop a satisfactory method for introducing heterologous genes into the genome of plastids. Furthermore, cellulase is expressed in all tissues which contain plastids which can be undesirable.

10 For production of ligninases to use in degrading lignins, the ligninases of choice are from the white-rot fungus *Phanerochaete chrysosporium*. One of the major lignin-degrading, extracellular enzymes produced by *P. chrysosporium* is lignin peroxidase (LIP). 15 Potential applications of LIP include not only lignin degradation but also biopulping of wood and biodegradation of toxic environmental pollutants. To produce large quantities of LIP, the fungus can be grown in large reactors and the enzyme isolated from the extracellular fluids. However, the yields have been low and the process has not been cost-effective. Production 20 of recombinant LIP in *E. coli*, in the fungus *Trichoderma reesei*, and baculovirus have been largely unsuccessful. Heterologous expression of lignin-degrading manganese 25 peroxidase in alfalfa plants has been reported; however, the transgenic plants had reduced growth and expression of the enzyme was poor (Austin *et al.*, *Euphytica* 85: 381-393 (1995)).

30 Although difficult to sufficiently and cheaply produce ligninases in non-plant systems, ligninases have evoked worldwide interest because of their potential in degrading a variety of toxic xenobiotic compounds such as PCBs and benzo(a)pyrenes in the environment (Yadav *et al.*, *Appl. Environ. Microbiol.* 61: 2560-2565 (1995); 35 Reddy, *Curr. Opin. Biotechnol.* 6: 320-328 (1995); Yadav

et al., Appl. Environ. Microbiol. 61: 677-680 (1994)).

Therefore, a need remains for an economical method for making transgenic crop plants wherein the ligninase and cellulase genes are incorporated into the plant genome but wherein the ligninase and cellulase expression are restricted to particular plant tissues, e.g., the leaves, and the ligninase and cellulase products are directed to a plant organelle wherein it accumulates without damaging the transgenic plant.

SUMMARY OF THE INVENTION

The present invention provides transgenic plants which after harvest degrade the lignin and cellulose therein to fermentable sugars which can further be fermented to ethanol or other products. In particular, the transgenic plants comprise ligninase and cellulase genes from microbes operably linked to a DNA encoding a signal peptide which targets the fusion polypeptide produced therefrom to an organelle of the plant, in particular the chloroplasts. When the transgenic plants are harvested, the plants are ground to release the ligninase and cellulase which then degrade the lignin and cellulose of the transgenic plants to produce the fermentable sugars.

Therefore, the present invention provides a transgenic plant which degrades lignocellulose when the transgenic plant is ground to produce a plant material comprising (a) at least one DNA encoding a cellulase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the cellulase to an organelle of the transgenic plant; and (b) at least one DNA encoding a ligninase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the ligninase to the organelle of the transgenic

plant, wherein the transgenic plant degrades the lignocellulose when ground to produce the plant material.

Further, the present invention provides a transgenic plant which degrades lignins when the transgenic plant is ground to produce a plant material comprising at least one DNA encoding a ligninase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the ligninase to an organelle of the transgenic plant wherein the transgenic plant degrades the lignins when ground to produce the plant material.

Further still, the present invention provides a transgenic plant which degrades cellulose when the transgenic plant is ground to produce a plant material comprising at least one DNA encoding a cellulase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the cellulase to an organelle of the transgenic plant wherein the transgenic plant degrades the cellulose when ground to produce the plant material.

The present invention also provides a method for producing a transgenic plant which degrades lignocellulose when the transgenic plant is ground to produce a plant material.

In one embodiment, the method comprises (a) providing a first transgenic plant which includes a DNA encoding a cellulase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the cellulase to an organelle of the transgenic plant and a second transgenic plant which includes a DNA encoding a ligninase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the ligninase to the organelle of the transgenic plant; and

5 (b) mating by sexual fertilization the first and the second transgenic plants to produce a third transgenic plant which includes the first DNA encoding the cellulase and the second DNA encoding the ligninase, wherein the transgenic plant degrades the lignocellulose when ground to produce the plant material.

10 In a further embodiment of the above method, the progeny of the third transgenic plant are mated by sexual fertilization to a transgenic plant selected from the group consisting of the first, second, and third transgenic plants to produce a transgenic plant comprising multiples of genes encoding cellulases and ligninases.

15 The present invention further provides a method for converting lignocellulose in a plant material to fermentable sugars.

20 In one embodiment, the method comprises (a) providing a transgenic plant which includes at least one DNA encoding a cellulase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the cellulase to an organelle of the transgenic plant and at least one DNA encoding a ligninase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the ligninase to the organelle of the transgenic plant; (b) growing the transgenic plant for a time sufficient for the transgenic plant to accumulate a sufficient amount of the cellulase and the ligninase in the organelle of the transgenic plant; (c) harvesting the transgenic plant which has accumulated the cellulase and ligninase in the organelle of the transgenic plant; (d) grinding the transgenic plant for a time sufficient to produce the plant material wherein the cellulase and ligninase produced by the transgenic plant are released from the organelle of the transgenic plant; (e)

incubating the plant material for a time sufficient for the cellulase and ligninase in the plant material to produce the fermentable sugars from the lignocellulose in the plant material; and (f) extracting the fermentable sugars produced from the lignocellulose by the cellulase and the ligninase from the plant material.

In another embodiment, the method comprises (a) providing a transgenic plant which includes at least one DNA encoding a cellulase which is operably linked to a nucleotide sequence encoding a signal peptide wherein the signal peptide directs the cellulase to an organelle of the transgenic plant; (b) growing the transgenic plant for a time sufficient for the transgenic plant to accumulate a sufficient amount of the cellulase in the organelle of the transgenic plant; (c) harvesting the transgenic plant which has accumulated the cellulase in the organelle of the transgenic plant; (d) grinding the transgenic plant for a time sufficient to produce a plant material wherein the cellulase is released from the organelle in the transgenic plant; (e) mixing the plant material with a fungus that produces a ligninase; (f) incubating the transgenic plant material with the fungus for a time sufficient for the cellulase released from the transgenic plant and the ligninase provided by the fungus to degrade the lignocellulose in the plant material to produce the fermentable sugars; and (g) extracting the fermentable sugars produced from the lignocellulose in the plant material.

In a further embodiment of either one of the above embodiments of the method for converting lignocellulose in a plant material to fermentable sugars, the plant material further includes a plant material made from a non-transgenic plant.

In a further still embodiment of the either one of the above methods, the fermentable sugars are

fermented to ethanol.

In a further embodiment of the transgenic plant for any one of the aforementioned embodiments of the present invention wherein the transgenic plant expresses cellulase and ligninase or the cellulase without the ligninase, the DNA encoding the cellulase is from an organism selected from the group consisting of *Trichoderma reesei*, *Acidothermus cellulolyticus*, *Streptococcus salivarius*, *Actinomyces naeslundi*, and *Thermomonospora fusca*.

In a further still embodiment of the above transgenic plant, the DNA encoding the cellulase is selected from the group consisting of an *el* gene from *Acidothermus cellulolyticus*, a *cbh1* gene from *Trichoderma reesei*, a dextranase gene from *Streptococcus salivarius*, and a beta-glucosidase gene from *Actinomyces naeslundi*.

In a further still embodiment of the above transgenic plant, the *el* gene comprises the nucleotide sequence set forth in SEQ ID NO:4, the *cbh1* gene comprises the nucleotide sequence set forth in SEQ ID NO:10, the dextranase gene comprises the nucleotide sequence set forth in SEQ ID NO:8, and the beta-glucosidase gene comprises the nucleotide sequence set forth in SEQ ID NO:6.

In a further still embodiment for any one of the aforementioned transgenic plants of the present invention wherein the transgenic plant expresses cellulase and ligninase or the ligninase without the cellulase, the DNA encoding the ligninase is from *Phanerochaete chrysosporium*.

In a further still embodiment of the above transgenic plant, the ligninase is *ckg4* comprising the nucleotide sequence set forth in SEQ ID NO:11 or *ckg5* comprising the nucleotide sequence set forth in SEQ ID

NO:13.

In a further still embodiment for any one of the aforementioned transgenic plants of the present invention, the DNA encoding the cellulase and the DNA 5 encoding the ligninase are each operably linked to a leaf-specific promoter. In a particular embodiment, the leaf-specific promoter is a promoter for *rbcS*.

In a further still embodiment of the above 10 transgenic plant, the nucleotide sequence encoding the signal peptide encodes a signal peptide of *rbcS*.

In a further still embodiment of the above transgenic plant, the *rbcS* comprises the nucleotide sequence set forth in SEQ ID NO:1.

In a further still embodiment for any one of the aforementioned transgenic plants of the present invention, the transgenic plant is selected from the group consisting of maize, wheat, barley, rye, hops, hemp, rice, potato, soybean, sorghum, sugarcane, clover, tobacco, alfalfa, *arabidopsis*, coniferous tree, and 15 deciduous tree.

In a further still embodiment for any one of the aforementioned transgenic plants of the present invention, the DNAs encoding the cellulase, ligninase, or both are stably integrated into nuclear or plastid 20 DNA of the transgenic plant.

In a further still embodiment for any one of the aforementioned transgenic plants of the present invention, the transgenic plant further includes a DNA 25 encoding a selectable marker operably linked to a constitutive promoter.

In a further still embodiment of the above transgenic plant, the DNA encoding the selectable marker provides the transgenic plant with resistance to an antibiotic, an herbicide, or to environmental stress.

35 In a further still embodiment of the above

transgenic plant, the DNA encoding resistance to the herbicide is a DNA encoding phosphinothricin acetyl transferase which confers resistance to the herbicide phosphinothricin.

5 In a further still embodiment for any one of the aforementioned transgenic plants of the present invention, the organelle of the transgenic plant is selected from the group consisting of nucleus, microbody, endoplasmic reticulum, endosome, vacuole, 10 mitochondria, chloroplast, or plastid.

15 In a further still embodiment of the above transgenic plant, the organelle of the transgenic plant is the chloroplast.

20 In a further still embodiment of the transgenic plants which degrade lignocellulose or cellulose, the lignocellulose is degraded to fermentable sugars which can then be fermented to ethanol.

OBJECTS

25 It is an object of the present invention to provide transgenic plants which degrade lignocellulose to fermentable sugars, methods for making the transgenic plants which degrade lignocellulose, and methods for using the transgenic plants to degrade lignocellulose to fermentable sugars.

30 These and other objects of the present invention will become increasingly apparent with reference to the following drawings and preferred embodiments.

DESCRIPTION OF THE DRAWINGS

35 Figure 1 is a diagram of a plasmid containing a heterologous gene expression cassette containing *cbl1* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide and a heterologous gene

expression cassette containing the *bar* gene operably linked to the *Act1* promoter. *rbcSP* is the *rbcS* gene promoter, *SP* is DNA encoding the *rbcS* signal peptide, *pin3'* is the 3' untranslated region of the potato inhibitor II-chloramphenicol acetyltransferase gene, *Act1* is the promoter for the *act1* gene, and *nos* is the 3' untranslated region of the *Agrobacterium* nopaline synthase gene.

Figure 2 is a diagram of a plasmid containing a heterologous gene expression cassette containing *el* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide and a heterologous gene expression cassette containing the *bar* gene operably linked to the *Act1* promoter. The terms in the diagram are as in Figure 1.

Figure 3 is a diagram of a heterologous gene expression cassette containing the *bar* gene in plasmid pDM302. *Act1* is the promoter for the *act1* gene and *nos* is the 3' untranslated region of the *Agrobacterium* nopaline synthase gene.

Figure 4 is a diagram of plasmid pSMF13 which is plasmid pSK containing a heterologous gene expression cassette containing *cbh1* operably linked to the *rbcS* promoter. The terms in the diagram are as in Figure 1.

Figure 5 is a diagram of plasmid pMSF14 which is plasmid pSK containing a heterologous gene expression cassette containing *cbh1* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide. The terms in the diagram are as in Figure 1.

Figure 6 is a diagram of plasmid pMSF15 which is plasmid pBI221 containing a heterologous gene expression cassette containing *syn-cbh1* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide. The terms in the diagram are as in Figure 1.

Figure 7 is a diagram of plasmid pTZA8 which is plasmid pBI121 containing a heterologous gene expression cassette containing *e1* operably linked to the CaMV35S promoter and DNA encoding the SSU signal peptide. SSU is the glycine max (soybean) *rbcS* signal peptide. CaMV35S is the cauliflower mosaic virus 35S promoter. The remainder of the terms are as in the diagram are as in Figure 1.

Figure 8 is a diagram of plasmid pZA9 which is plasmid pBI121 containing a heterologous gene expression cassette containing *e1* operably linked to the CaMV35S promoter and DNA encoding the VSP signal peptide. VSP is the soybean vegetative storage protein beta-leader sequences. The remainder of the terms in the diagram are as in Figure 7.

Figure 9 is a diagram of plasmid pZA10 which is plasmid pBI121 containing a heterologous gene expression cassette containing *e1* operably linked to the CaMV35S promoter. The remainder of the terms in the diagram are as in Figure 7.

Figure 10 is a diagram of a plasmid containing a heterologous gene expression cassette containing *ckg4* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide and a gene expression cassette containing the *bar* gene operably linked to the *Act1* promoter. The remainder of the terms in the diagram are as in Figure 1.

Figure 11 is a diagram of a plasmid containing a heterologous gene expression cassette containing *ckg5* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide and a gene expression cassette containing the *bar* gene operably linked to the *Act1* promoter. The remainder of the terms in the diagram are as in Figure 1.

Figure 12 is a diagram of plasmid pSMF18 containing a heterologous gene expression cassette containing *ckg4* operably linked to the *rbcS* promoter. The remainder of the terms in the diagram are as in Figure 1.

5

Figure 13 is a diagram of plasmid pSMF19 containing a heterologous gene expression cassette containing *ckg5* operably linked to the *rbcS* promoter. The remainder of the terms in the diagram are as in Figure 1.

10

Figure 14 is a diagram of plasmid pSMF16 containing a heterologous gene expression cassette containing *ckg4* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide. The remainder of the terms in the diagram are as in Figure 1.

15

Figure 15 is a diagram of plasmid pSMF17 containing a heterologous gene expression cassette containing *ckg5* operably linked to the *rbcS* promoter and DNA encoding the *rbcS* signal peptide. The remainder of the terms in the diagram are as in Figure 1.

20

DESCRIPTION OF PREFERRED EMBODIMENTS

All patents, patent applications, government publications, government regulations, and literature references cited in this specification are hereby incorporated herein by reference in their entirety. In case of conflict, the present description, including definitions, will control.

25

The term "cellulase" is used herein as a generic term that includes endoglucanases such as the EI beta-1,4-endoglucanase precursor gene (el) of *Acidothermus cellulolyticus* and exoglucanases such as the cellobiohydrolase gene (cbh1) of *Trichoderma reesei* (also classified by some as *Trichoderma*

30

longibrachiatum), the dextranase gene of *Streptococcus salivarius* encoding the 1,6-alpha-glucanhydrolase gene, and the beta-glucosidase gene from *Actinomyces naeslundi*. Endoglucanases randomly cleave cellulose chains into smaller units. Exoglucanases include cellobiohydrolases, which liberate glucose dimers (cellobiose) from the ends of cellulose chains; glucanhydrolases, which liberate glucose monomers from the ends of cellulose chains; and, beta-glucosidases, which liberate D-glucose from cellobiose dimers and soluble celldextrins. When all four of the above enzymes are combined, they work synergistically to rapidly decrystallize and hydrolyze cellulose to fermentable sugars.

The term "lignin" is used herein as a generic term that includes both lignins and lignocelluloses.

The term "ligninase" is used herein as a generic term that includes all varieties of enzymes which degrade lignins such as the lignin peroxidase gene of *Phanerochaete chrysosporium*.

A variety of fungi and bacteria produce ligninase and cellulase enzymes, and based on evolutionary pressures, these fungi are able to degrade lignin or cellulose and hemicellulose of plant residues in the soil. In the laboratory, cellulases have been used to hydrolyze or convert cellulose and hemicellulose into mixtures of simple sugars that can be used in fermentation to produce a wide variety of useful chemical and fuel products, including but not limited to, ethanol, lactic acid, and 1,3-propanediol, which is an important molecular building block in the production of environmentally-friendly plastics.

The biodegradation of lignin, which comprises 20-30% of the dry mass of woody plants, is of great economic importance because this process is believed to

be an important rate-limiting step in the earth's carbon cycle. Furthermore, there is considerable potential for the transformation of lignin into aromatic chemical feedstock. Also, delignification of lignocellulosic feeds has been shown to increase their digestibility by cattle by about 30%, therefore, contributing to enhanced cost effectiveness for producing milk and meat. Moreover, research on lignin biodegradation has important implications in biopulping and biobleaching in the paper industry.

The present invention provides transgenic plants which produce ligninases, cellulases, or both in the leaves and straw/stalks of the plant. While the transgenic plant can be any plant which is practical for commercial production, it is preferable that the transgenic plants be constructed from plants which are produced in large quantities and which after processing produce a substantial amount of leaves and stalks as a byproduct. Therefore, it is desirable that the transgenic plant be constructed from plants including, but not limited to, maize, wheat, barley, rye, hops, hemp, rice, potato, soybean, sorghum, sugarcane, clover, tobacco, alfalfa, arabidopsis, coniferous trees, and deciduous trees. Most preferably, the transgenic plant is constructed from maize.

Maize is a preferred plant because it is a major crop in the United States; approximately 60 million acres of maize are produced per year. Further, there is already a large industry built around the processing of maize grain to industrial products, which includes the production of over 1.2 billion gallons of fuel ethanol per year. Thus, fermentable sugars produced by the hydrolysis of maize stalks and leaves according to the present invention can be utilized within the large existing maize refining infrastructure.

Leaves and stalks from transgenic maize made according to the present invention can be made available to this refining infrastructure in large quantities, about tens of millions of tons annually) at a current cost of about 5 30 dollars per ton. This cost is about one quarter of the cost of maize grain which further enhances the value of the present invention for the economical production of a wide variety of industrial products from the residue of transgenic plants made according to the 10 present invention. Furthermore, maize is preferred because it is a C-4 monocot that has very large chloroplasts. The large chloroplasts enables the chloroplasts of the transgenic maize of the present invention to accumulate higher levels of ligninases and 15 cellulases than could be accumulated in the chloroplasts of other transgenic plants, e.g., C-3 dicots and monocots. Therefore, transgenic maize of the present invention is a particularly useful source of ligninases and cellulases.

20 Thus, the transgenic plants of the present invention provide a plentiful, inexpensive source of fungal or bacterial ligninases and cellulases which can be used to degrade lignins and cellulose in plants to fermentable sugars for production of ethanol or for 25 other uses which require ligninases and cellulases such as pre-treating silage to increase the energy value of lignocellulosic feeds for cows and other ruminant animals, pre-treating lignocellulosic biomass for 30 fermentative conversion to fuels and industrial chemicals, and biodegradation of chloroaromatic environmental pollutants. Because the transgenic plants of the present invention produce the ligninases, cellulases, or both therein, the external addition of 35 ligninases and cellulases for degradation of the plant material is no longer necessary. Therefore, the present

invention enables the plant biomass, which is destined to become ethanol or other products, to serve as the source of ligninase and cellulase. Furthermore, the plant material from the transgenic plants of the present invention can be mixed with non-transgenic plant material. The ligninases, cellulases, or both produced by the transgenic plants will degrade the lignin and cellulose of all the plant material, including the non-transgenic plant material. Thus, ligninase and cellulase degradation of plant material can be carried out more economically.

The transgenic plants of the present invention comprise one or more heterologous gene expression cassettes containing DNA encoding at least one fungal or bacterial ligninase, cellulase, or both inserted into the plant's nuclear genome. The preferred cellulase is encoded by a DNA from the microorganism *Acidothermus cellulolyticus*, *Thermomonospora fusca*, and *Trichoderma reesei* (*Trichoderma longibrachiatum*). Other microorganisms which produce cellulases suitable for the present invention include *Zymomonas mobilis*, *Acidothermus cellulolyticus*, *Clostridium thermocellum*, *Eiwinia chrysanthemi*, *Xanthomonas campestris*, Alkalophilic *Bacillus* sp., *Cellulomonas fimi*, wheat straw mushroom (*Agaricus bisporus*), *Ruminococcus flavefaciens*, *Ruminococcus albus*, *Fibrobacter succinogenes*, and *Butyrivibrio fibrisolvens*.

The preferred ligninase is lignin peroxidase (LIP) encoded in DNA from *Phanerochaete chrysosporium* or *Phlebia radiata*. One of the major lignin-degrading, extracellular enzymes produced by *P. chrysosporium* is LIP. The LIPs are glycosylated heme proteins (MW 38 to 46 kDa) which are dependent on hydrogen peroxide for activity and catalyze the oxidative cleavage of lignin

polymer. At least six heme proteins (H1, H2, H6, H7, H8, and H10) with LIP activity have been identified in P. chrysosporium strain BKMF-1767 of which isozymes H2, H6, H8, and H10 are the major LIPs in both static and 5 agitated cultures of *P. chrysosporium*. However, other fungi which produce ligninases suitable for use in the present invention include *Bjerkandera adusta*, *Trametes hirsuta*, *Plebia radiata*, *Pleurotus* spp., *Stropharia aurantiaca*, *Hypholoma fasciculare*, *Trametes versicolor*, 10 *Gymnopilus penetrans*, *Stereum hirsutum*, *Mycena haematopus*, and *Armillaria mellea*.

In the present invention, the transgenic plant 15 comprises a DNA encoding one or more cellulase fusion proteins wherein the DNA encoding the cellulases are operably linked to a DNA encoding a signal peptide which directs the cellulase fusion protein to a plant organelle such as the nucleus, a microbody (e.g., a peroxisome, or specialized version thereof, such as a glyoxysome), an endoplasmic reticulum, an endosome, a 20 vacuole, a mitochondria, a chloroplast, or a plastid. By sequestering the cellulase fusion proteins in the plant organelle, the cellulase fusion protein is prevented from leaking outside the cytoplasm to harm the plant by degrading the cellulose in the plant's cell 25 wall while the plant is being cultivated. In particular embodiments of the present invention, the gene encoding the cellulase is modified by replacing the amino acid codons that encode the leader region of the cellulase with amino acid codons that encode the signal peptide.

30 In a preferred embodiment of the invention, the amino acid codons that encode the signal peptide that directs the protein to which it is attached to the plant organelle, the chloroplasts, are the nucleotide codons that encode the rice rubisco synthase gene (*rbcS*) 35 small subunit signal peptide (*rbcSSP*). The nucleotide

sequence of the *rbcS* is set forth in SEQ ID NO:1 (GenBank Accession No. X07515). The 47 amino acid signal peptide of the *rbcS* protein has the amino acid sequence MAPPS VMASS ATIVA PFQGS SPPPA CRRPP SELQL RQRQH 5 GGRIR CM (SEQ ID NO:2). The *rbcS* SP directs proteins to which it is operably linked to the chloroplasts of the transgenic plant. Therefore, in the preferred embodiment of the present invention, the transgenic plant comprises a DNA encoding the cellulase operably linked with a DNA encoding the *rbcS* SP to produce the cellulase fusion protein. The *rbcS* SP directs the cellulase fusion protein to the chloroplasts. Thus, the cellulase fusion protein produced by the transgenic plant accumulates in the chloroplasts of the transgenic 10 plant which protects the transgenic plant from degradation by the cellulase fusion protein while it is being cultivated. Alternatively, the DNA encoding the cellulase is modified at its 3'end to encode a transit peptide such as the peptide RAVARL (SEQ ID NO:3), which 15 targets the ligninase fusion protein to the peroxisomes (U.S. Patent 6,103,956 to Srienc *et al.*). Preferably, the leader region of the cellulase is also removed. In any one of the above embodiments, the cellulase can be 20 further modified to include a GC content that approximates the GC content of the genomic DNA of the plant by methods well known in the art.

In a preferred embodiment, the cellulase comprising the cellulase fusion protein is encoded by the EI beta-1,4-endoglucanase precursor gene (*e1*) of *Acidothermus cellulolyticus*, the cellobiohydrolase gene (*cbh1*) of *Trichoderma reesei* (*Trichoderma longibrachiatum*), the beta-glucosidase gene from *Actinomyces naeslundi*, or the glucanhydrolase (dextranase) gene from *Streptococcus salivarius*. The 25 30

nucleotide sequence of the e1 DNA is set forth in SEQ ID NO:4 (GenBank Accession No. U33212), which encodes the cellulase with the amino acid sequence set forth in SEQ ID NO:5. SEQ ID NO:6 provides the nucleotide sequence of the beta-glucosidase gene from *Actinomyces naeslundi* (GenBank Accession No. AY029505), which encodes the beta-glucosidase with the amino acid sequence set forth in SEQ ID NO:7. SEQ ID NO:8 provides the nucleotide sequence of the dextranase gene from *Streptococcus salivarius* (GenBank Accession No. D29644), which encodes a glucanhydrolase with the amino acid sequence set forth in SEQ ID NO:9. The nucleotide sequence of *cbh1* is set forth in SEQ ID NO:10 (GenBank Accession No. E00389), which encodes the cellulase that includes the joined exons from positions 210 to 261, 738 to 1434, and 1498-1881.

In the present invention, the transgenic plant comprises a DNA encoding one or more ligninase fusion proteins wherein a DNA encoding the ligninase is operably linked to a DNA encoding a signal peptide which directs the ligninase fusion protein to a plant organelle. By sequestering the ligninase fusion proteins in the plant organelles, the modified ligninase is prevented from leaking outside the cytoplasm to harm the plant by degrading the ligninase in the plant's cell wall while the plant is being cultivated. In particular embodiments of the present invention, the leader sequence of the gene encoding the ligninase is modified by replacing the amino acid codons that encode the leader region of the ligninase with amino acid codons that encode the signal peptide.

In a preferred embodiment of the invention, the amino acid codons that encode the signal peptide are the amino acid codons which encode the rice rubisco synthase gene (*rbcS*) small subunit signal peptide

(*rbcSSP*). The nucleotide sequence of the *rbcS* is set forth in SEQ ID NO:1 (GenBank Accession No. X07515). The 47 amino acid signal peptide of the *rbcS* protein has the amino acid sequence MAPPS VMASS ATIVA PFQGS SPPPA 5 CRRPP SELQL RQRQH GGRIR CM (SEQ ID NO:2). Therefore, in the preferred embodiment of the present invention, the transgenic plant comprises a DNA encoding the ligninase operably linked to a DNA encoding the *rbcS* SP. The *rbcS* SP directs the ligninase fusion protein to the 10 chloroplasts. Thus, the ligninase fusion protein produced by the transgenic plant accumulates in the chloroplasts of the transgenic plant which protects the transgenic plant from degradation by the ligninase fusion protein while it is being cultivated. 15 Alternatively, the DNA encoding the ligninase is modified at its 3'end to encode a transit peptide such as the peptide RAVARL (SEQ ID NO:3). Optionally, the leader region of the ligninase is also removed. In any one of the above embodiments, the ligninase can be 20 further modified to include a GC content that approximates the GC content of the genomic DNA of the plant by methods well known in the art.

In a preferred embodiment of the invention, the ligninase comprising the ligninase fusion protein is 25 encoded by the lignin peroxidase gene (LIP) genes *ckg4* (H2) and *ckg5* (H10) of *Phanerochaete crysosporium* (de Boer et al., Gene 6: 93-102 (1987), Corrigendum in Gene 69: 369 (1988)). The nucleotide sequence of the *ckg4* gene is set forth in SEQ ID NO:11 (GenBank Accession No. 30 M18743), which encodes the amino acid with the sequence set forth in SEQ ID NO:12. The nucleotide sequence of the *ckg5* gene is set forth in SEQ ID NO:13 (GenBank Accession No. M18794), which encodes the amino acid with the sequence set forth in SEQ ID NO:14.

35 In the present invention, transcription and,

therefore, expression of the ligninase and cellulase fusion proteins are effected by a promoter that is active in a particular tissue of the plant, e.g., a promoter that is active primarily in the leaves of a plant. A leaf-specific promoter that is preferred for transcription (expression at the RNA level) is the rice rubisco synthase gene promoter (*rbcSP*), which has the nucleotide sequence prior to the *rbcS* gene coding region included in SEQ ID NO:1. In some embodiments of the present invention, it is desirable to relegate transcription of the heterologous gene expression cassette to the seeds using a seed-specific promoter. Seed-specific promoters that are suitable include, but are not limited to, the seed-specific promoters such as the maize 19 kDa zein (cZ19B1) promoter, the maize cytokinin-induced message (*Cim1*) promoter, and the maize myo-inositol-1-phosphate synthase (*milps*) promoter, which are disclosed in U.S. Patent 6,225,529 to Lappegard *et al.* Therefore, in the heterologous gene expression cassettes, the nucleotide sequence comprising *rbcS* promoters are operably linked to the nucleotide sequences encoding the ligninase and cellulase fusion proteins. Thus, in a transgenic plant of the present invention, transcription of the ligninase and cellulase fusion proteins occurs primarily in the leaves of the plant, and because the ligninase and cellulase fusion proteins each has a signal peptide that directs its transport to plastids, the ligninase and cellulase fusion proteins accumulate in the plastids.

In the preferred embodiment of the present invention, the 3' ends of the nucleotide sequence encoding the above ligninase and cellulase fusion proteins are operably linked to a 3' noncoding sequence wherein the noncoding sequence contains a poly(A) cleavage/addition site and other regulatory sequences

which enables the RNA transcribed therefrom to be properly processed and polyadenylated which in turn affects stability, transport and translation of the RNA transcribed therefrom in the plant cell. Examples of 3' 5 noncoding sequences include the 3' noncoding sequence from the potato protease inhibitor II gene, which includes nucleotides 871 to 1241 of SEQ ID NO: 15 (GenBank Accession No. M15186) and the 3' noncoding sequence from the *Agrobacterium* nopaline synthase gene, 10 which includes nucleotides 2001 to 2521 of SEQ ID NO:16 (GenBank Accession No.V00087 J01541).

The above heterologous gene expression cassettes can be constructed using conventional molecular biology cloning methods. In a particularly convenient method, PCR is used to produce the nucleotide fragments for constructing the gene expression cassettes. By using the appropriate PCR primers, the precise nucleotide regions of the above DNAs can be amplified to produce nucleotide fragments for cloning. 15 By further including in the PCR primers restriction enzyme cleavage sites which are most convenient for assembling the heterogenous gene expression cassettes (e.g., restriction enzyme sites that are not in the nucleotide fragments to be cloned), the amplified 20 nucleotide fragments are flanked with the convenient restriction enzyme cleavage sites for assembling the nucleotide fragments into heterogenous gene expression cassettes. The amplified nucleotide fragments are 25 assembled into the heterogeneous gene expression cassettes using conventional molecular biology methods. Based upon the nucleotide sequences provided herein, how 30 to construct the heterogenous gene expression cassettes using conventional molecular biology methods with or without PCR would be readily apparent to one skilled in 35 the art.

5 In a further embodiment of the present invention, the transgenic plant comprises more than one heterogeneous gene expression cassette. For example, the transgenic plant comprises a first cassette which
10 contains a DNA encoding a ligninase fusion protein, and one or more cassettes each containing a DNA encoding a particular cellulase fusion protein. Preferably, both the ligninase and cellulase fusion proteins comprise amino acids of a signal peptide which directs the fusion proteins to plant organelles. In a preferred embodiment, the signal peptide for each is the *rbcS* SP or the SKL motif.

15 In a further still embodiment, the transgenic plant comprises DNA encoding the ligninase fusion protein such as the *ckg4* or *ckg5* LIP, an endoglucanase fusion protein such as the *e1* fusion protein, and a cellobiohydrolase fusion protein such as the *cbh1* fusion protein. In a further still embodiment, the transgenic plant comprises DNA encoding the ligninase fusion protein such as the *ckg4* or *ckg5* LIP, an endoglucanase fusion protein such as the *e1* fusion protein, a cellobiohydrolase fusion protein such as the *cbh1* fusion protein, a beta-glucosidase, and a glucanhydrolase. Preferably, both the ligninase and cellulase fusion
20 proteins comprise amino acids of a signal peptide which directs the fusion proteins to plant organelles. In a preferred embodiment, the signal peptide for each is the *rbcS* SP or the SKL motif.

25

30 To make the transgenic plants of the present invention, plant material such as meristem primordia tissue is transformed with plasmids, each containing a particular heterogenous gene expression cassette using the Biolistic bombardment method as described in Example 5 and in U.S. Patent No. 5,767,368 to Zhong et al.
35 Further examples of the Biolistic bombardment method are

disclosed in U. S. application Ser. No. 08/036,056 and U. S. Patent No. 5,736,369 to Bowen *et al.* Each heterogenous gene expression cassette is separately introduced into a plant tissue and the transformed tissue propagated to produce a transgenic plant that contains the particular heterogenous gene expression cassette. Thus, the result is a transgenic plant containing the heterogenous gene expression cassette expressing a ligninase such as *ckg4* or *ckg5*, a transgenic plant containing a heterogenous gene expression cassette expressing endoglucanase such as *e1*, a transgenic plant containing a heterogenous gene expression cassette expressing a cellobiohydrolase such as *cbh1*, a transgenic plant containing a heterogenous gene expression cassette expressing an exoglucanase such as beta-glucosidase, and a transgenic plant containing a heterogenous gene expression cassette expressing an exoglucanase such as glucanhydrolase.

Alternatively, transformation of corn plants can be achieved using electroporation or bacterial mediated transformation using a bacterium such as *Agrobacterium tumefaciens* to mediate the transformation of corn root tissues (see Valvekens *et al.* Proc. Nat'l. Acad. Sci. USA. 85: 5536-5540 (1988)) or meristem primordia.

In a preferred embodiment of the present invention, the transgenic plant comprises one or more ligninase fusion proteins and one or more cellulase fusion proteins. Construction of the preferred transgenic plant comprises making first generation transgenic plants as above, each comprising a ligninase fusion protein, and transgenic plants as above, each comprising a cellulase fusion protein. After each first generation transgenic plant has been constructed, progeny from each of the first generation transgenic

plants are cross-bred by sexual fertilization to produce second generation transgenic plants comprising various combinations of both the ligninase fusion protein and the cellulase fusion protein.

5 For example, various combinations of progeny from the first generation transgenic plants are cross-bred to produce second generation transgenic plants that contain *ckg4* and *cbh1*, *e1*, beta-glucosidase, or *ckg5*; second generation transgenic plants that contain *ckg5* and *cbh1*, *e1*, or beta-glucosidase; second generation transgenic plants that contain *e1* or beta glucosidase, and a second generation transgenic plant that contains *e1* and beta-glucosidase.

10 15 Progeny of the second generation transgenic plants are cross-bred by sexual fertilization among themselves or with first generation transgenic plants to produce third generation transgenic plants that contain one or more ligninases, one or more cellulases, or combinations thereof.

20 25 For example, cross-breeding a second generation transgenic plant containing *ckg4* and *cbh1* with a second generation transgenic plant containing *e1* and beta-glucosidase produces a third generation transgenic plant containing *ckg4*, *cbh1*, *e1*, and beta-glucosidase. The third generation transgenic plant can be cross-bred with a first generation transgenic plant containing *ckg5* to produce a fourth generation transgenic plant containing *ckg4*, *ckg5*, *cbh1*, *e1*, and beta-glucosidase.

30 35 It will be readily apparent to one skilled in the art that other transgenic plants with various combinations of ligninases and cellulases can be made by cross-breeding progeny from particular transgenic plants. Zhang et al, Theor. Appl. Genet. 92: 752-761, (1996), Zhong et al, Plant Physiol. 110: 1097-1107,

(1996);, and Zhong et al, *Planta*, 187: 483-489, (1992) provide methods for making transgenic plants by sexual fertilization.

5 Alternatively, plant material is transformed as above with a plasmid containing a heterologous gene expression cassette encoding the ligninase fusion protein. The transgenic plant is recovered from the progeny of the transformed plant material. Next, plant material from the transgenic plant is transformed with 10 a second plasmid containing a heterologous gene expression cassette encoding the cellulase fusion protein and a second selectable marker. The transgenic plant is recovered from the progeny of the transformed plant material. It will be readily apparent to one skilled in the art that transgenic plants containing any 15 combination of ligninases and cellulases can be made by the above method.

20 In a preferred embodiment, the above heterologous gene expression cassettes further include therein nucleotide sequences that encode one or more selectable markers which enable selection and identification of transgenic plants that express the modified cellulase of the present invention. Preferably, the selectable markers confers additional 25 benefits to the transgenic plant such as herbicide resistance, insect resistance, and/or resistance to environmental stress.

30 Alternatively, the above transformations are performed by co-transforming the plant material with a first plasmid containing a heterologous gene expression cassette encoding a selectable marker and a second plasmid containing a heterologous gene expression cassette encoding a ligninase or cellulase fusion protein. The advantage of using a separate plasmid is 35 that after transformation, the selectable marker can be

removed from the transgenic plant by segregation, which enables the selection method for recovering the transgenic plant to be used for recovering transgenic plants in subsequent transformations with the first
5 transgenic plant.

Examples of preferred markers that provide resistance to herbicides include, but are not limited to, the *bar* gene from *Streptomyces hygroscopicus* encoding phosphinothricin acetylase (PAT), which confers
10 resistance to the herbicide glufonsinate; mutant genes which encode resistance to imidazalinone or sulfonylurea such as genes encoding mutant form of the ALS and AHAS enzyme as described by Lee *et al.* EMBO J. 7: 1241 (1988) and Miki *et al.*, Theor. Appl. Genet. 80: 449 (1990), respectively, and in U.S. Patent No. 5,773,702 to Penner
15 *et al.*; genes which confer resistance to glycophosphate such as mutant forms of EPSP synthase and *aroA*; resistance to L-phosphinothricin such as the glutamine synthetase genes; resistance to glufosinate such as the phosphinothricin acetyl transferase (PAT and *bar*) gene;
20 and resistance to phenoxy propionic acids and cyclohexones such as the ACCase inhibitor-encoding genes (Marshall *et al.* Theor. Appl. Genet. 83: 435 (1992)). The above list of genes which can import
25 resistance to an herbicide is not inclusive and other genes not enumerated herein but which have the same effect as those above are within the scope of the present invention.

Examples of preferred genes which confer resistance to pests or disease include, but are not limited to, genes encoding a *Bacillus thuringiensis* protein such as the delta-endotoxin, which is disclosed
30 in U.S. Patent 6,100,456 to Sticklen *et al.*; genes encoding lectins, (Van Damme *et al.*, Plant Mol. Biol. 24: 825 (1994)); genes encoding vitamin-binding proteins
35

such as avidin and avidin homologs which can be used as larvicides against insect pests; genes encoding protease or amylase inhibitors, such as the rice cysteine proteinase inhibitor (Abe *et al.*, *J. Biol. Chem.* 262: 5 16793(1987)) and the tobacco proteinase inhibitor I (Hubb *et al.*, *Plant Mol. Biol.* 21: 985(1993)); genes encoding insect-specific hormones or pheromones such as ecdysteroid and juvenile hormone, and variants thereof, mimetics based thereon, or an antagonists or agonists thereof; genes encoding insect-specific peptides or neuropeptides which, upon expression, disrupts the physiology of the pest; genes encoding insect-specific venom such as that produced by a wasp, snake, etc.; genes encoding enzymes responsible for the accumulation of monoterpenes, sesquiterpenes, asteroid, hydroxamic acid, phenylpropanoid derivative or other non-protein molecule with insecticidal activity; genes encoding enzymes involved in the modification of a biologically active molecule (see U.S. Patent No. 5,539,095 to 10 Sticklen *et al.*, which discloses a chitinase that functions as an anti-fungal); genes encoding peptides which stimulate signal transduction; genes encoding hydrophobic moment peptides such as derivatives of *Tachyplesin* which inhibit fungal pathogens; genes 15 encoding a membrane permease, a channel former or channel blocker (for example cecropin-beta lytic peptide analog renders transgenic tobacco resistant to *Pseudomonas solanacearum*) (Jaynes *et al.* *Plant Sci.* 89: 43 (1993)); genes encoding a viral invasive protein or 20 complex toxin derived therefrom (viral accumulation of viral coat proteins in transformed cells of some transgenic plants impart resistance to infection by the virus the coat protein was derived as shown by Beachy *et al.* *Ann. Rev. Phytopathol.* 28: 451 (1990); genes 25 encoding an insect-specific antibody or antitoxin or a 30 35

virus-specific antibody (Tavladoraki *et al.* *Nature* 366: 469(1993)); and genes encoding a developmental-arrestive protein produced by a plant, pathogen or parasite which prevents disease. The above list of genes which can import resistance to disease or pests is not inclusive and other genes not enumerated herein but which have the same effect as those above are within the scope of the present invention.

Examples of genes which confer resistance to environmental stress include, but are not limited to, *mtld* and *HVA1*, which are genes that confer resistance to environmental stress factors; *rd29A* and *rd19B*, which are genes of *Arabidopsis thaliana* that encode hydrophilic proteins which are induced in response to dehydration, low temperature, salt stress, or exposure to abscisic acid and enable the plant to tolerate the stress (Yamaguchi-Shinozaki *et al.*, *Plant Cell* 6: 251-264 (1994)). Other genes contemplated can be found in U.S. Patents Nos. 5,296,462 and 5,356,816 to Thomashow. The above list of genes, which can import resistance to environmental stress, is not inclusive and other genes not enumerated herein but which have the same effect as those above are within the scope of the present invention.

Thus, it is within the scope of the present invention to provide transgenic plants which express one or more ligninase fusion proteins, one or more cellulase fusion proteins, and one or more of any combination of genes which confer resistance to an herbicide, pest, or environmental stress.

In particular embodiments of the present invention, the heterologous gene expression cassettes can further be flanked with DNA containing the matrix attachment region (MAR) sequence. While use of MAR in the present invention is optional, it can be used to

increase the expression level of transgenes, to get more reproducible results, and to lower the average copy number of the transgene (Allen et al., *The Plant Cell* 5: 603-613 (1993); Allen et al., *The Plant Cell* 8: 899-913 5 (1996); Mlynarova et al., *The Plant Cell* 8: 1589-1599 (1996)).

To degrade the lignocellulose in the leaves and stalks of the transgenic plants of the present invention, the transgenic plant is ground up to produce 10 a plant material using methods currently available in the art to disrupt a sufficient number of the plant organelles containing the ligninase and cellulase therein. The ligninase and cellulase degrade the lignocellulose of the transgenic plant into fermentable sugars, primarily glucose, and residual solids. The fermentable sugars are used to produce ethanol or other 15 products.

The transgenic plants can be processed to ethanol in an improvement on the separate saccharification and fermentation (SHF) method (Wilke et 20 al., *Biotechnol. Bioengin.* 6: 155-175 (1976)) or the simultaneous saccharification and fermentation (SSF) method disclosed in U.S. Patent 3,990,944 to Gauss et al. and U.S. Patent 3,990,945 to Huff et al. The SHF 25 and SSF methods require pre-treatment of the plant material feedstock with dilute acid to make the cellulose more accessible followed by enzymatic hydrolysis using exogenous cellulases to produce glucose from the cellulose, which is then fermented by yeast to ethanol. In some variations of the SHF or SSF methods, the plant material is pre-treated with heat or with both 30 heat and dilute acid to make the cellulose more accessible.

An SHF or SSF method that uses the transgenic 35 plant material of the present invention as the feedstock

is an improvement over the SHF or SSF method because the transgenic plant material contains its own cellulases and ligninases or cellulases. Therefore, exogenous ligninases and/or cellulases do not need to be added to the feedstock. Furthermore, because particular embodiments of the transgenic plant material produce ligninase, the need for pre-treatment of the plant material in those embodiments before enzymatic degradation is not necessary. In a further improvement over the SHF method, the transgenic plant material is mixed with non-transgenic plant material and the mixture processed to ethanol.

The following examples are intended to promote a further understanding of the present invention.

EXAMPLE 1

This example shows the construction of plasmids comprising a heterologous gene expression cassette comprising a DNA encoding a cellulase fusion protein and a heterologous gene expression cassette comprising a DNA encoding the *bar* gene (Table 1).

Table 1

	Construct	Plasmid features
1	<i>rbcSP/e1/pin 3'//Act1 P/bar/nos 3'</i>	<i>rbcSP</i> leaf-specific promoter driving cellulase cDNA of <i>A. cellulolyticus</i>
2	<i>rbcSP/cbh1/pin 3'//Act1 P/bar/nos 3'</i>	<i>rbcSP</i> leaf-specific promoter driving cellulase cDNA of <i>T. reesi</i>
3	<i>rbcSP/rbcS SP/e1/pin 3'//Act1 P/bar/nos 3'</i>	The <i>rbcS</i> SP targets cellulase of <i>A. cellulolyticus</i> into maize chloroplasts
4	<i>rbcSP/rbcS SP/cbh1/pin 3'//Act1 P/bar/nos 3'</i>	The <i>rbcS</i> SP targets cellulase of <i>T. reesi</i> into maize chloroplasts

30 Abbreviations:

The term "rbcSP" means the rice rubisco *rbcS* promoter region. The *rbcSP* is a leaf-specific promoter that limits transcription of *rbcS* to the leaves (Schaeffer and Sheen, *Plant Cell* 3: 997-1012 (1991)).
5 The nucleotide sequence for the *rbcS* promoter region is set forth in SEQ ID NO:1.

The term "e1" means the cDNA isolated from *Acidothermus cellulolyticus* which encodes the cellulase EI beta-1,4-endoglucanase precursor. The nucleotide sequence for the gene encoding e1 is set forth in SEQ ID NO:4. In this example, the codons for the 41 amino acid leader sequence (nucleotides 824 to 946 of SEQ ID NO:4) are removed.
10

The term "cbh1" means the cDNA isolated from *Trichoderma reesi* that encodes the cellulase cellobiohydrolase. The nucleotide sequence for the gene encoding cbh1 is set forth in SEQ ID NO:10. In this example, the codons for the 54 amino acid leader sequence (nucleotides 210 to 671 of SEQ ID NO:10) are removed.
15
20

The term "pin3'" means the potato protease inhibitor II-chloramphenicol acetyltransferase gene's 3' untranslated sequence which contains transcription termination signals (Thornburg et al., *Proc. Natl. Acad. Sci. USA* 84: 744-748 (1987)). The *pin3'* untranslated sequence includes nucleotides 882 to 1241 of the nucleotide sequence set forth in SEQ ID NO: 15.
25

The term "bar" means the phosphinothricin acetyl transferase gene (Thompson et al., *EMBO J.* 6: 2519-2523 (1987)). The *bar* gene is a selectable marker for herbicide resistance. The 5' end of *bar* is operably linked to the rice actin 1 gene promoter which has been shown to operable in maize (Zhong et al., *Plant Physiology* 110: 1097-1107 (1996); Zhang et al., *Theor.*

Appl. Genet. 92: 752-761 (1996); Zhang et al., Plant Science 116: 73-84 (1996)). The 3' end of *bar* is operably linked to the *nos* 3' untranslated sequences. The nucleotide sequence of the *bar* gene is set forth in SEQ ID NO:18 (GenBank Accession No. X05822), which encodes the *bar* having the amino acid sequence from nucleotides 160 to 711.

5 The term "Act1 P" means the rice *Act1* gene promoter which further includes the 5' intron region (McElroy et al., Mol. Gen. Genet. 231: 150-160 (1991)). The sequence of the *Act1* gene and its promoter is set forth in SEQ ID NO:19 (GenBank Accession No. X63830).

10 The term "nos3'" means the 3' untranslated sequence from the *Agrobacterium* nopaline synthase gene encoding nopaline synthase of the amino acid sequence as set forth in SEQ ID NO:17 which includes nucleotides 2002 to 2521 of SEQ ID NO:16 (GenBank Accession No. V00087 J01541). The *Nos3'* sequence contains transcription termination signals.

15 The term "rbcS SP" means the rice rubisco small subunit signal peptide which consists of 47 codons encoding the peptide with the amino acid sequence set forth in SEQ ID NO:2. The *rbcS* SP directs the translocation of the *rbcS* small subunit or any polypeptide to which it is attached to the chloroplasts (Loza-Tavera et al., Plant Physiol. 93: 541-548 (1990)).

20 Construct 1 contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the cellulase encoded by *el* to the cells of the leaves of the maize plant.

25 Construct 2 contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the cellulase encoded by *cbh1* to the cells of the leaves of the maize plant.

GPO 2000 O 500 2000 2000

Construct 3, which is shown in Figure 1, is like construct 1 except that DNA encoding the *rbcS* SP signal peptide is operably linked to the 5' end of the *el*, and construct 4, which is shown in Figure 2, is like construct 2 except that DNA encoding the *rbcS* SP signal peptide is operably linked to the 5' end of *cbh1*. Therefore, expression of cellulase from construct 3 or 4, which is limited to the cells of the leaves, directed to the chloroplasts in the cells. All of the above constructs are adjacent to a heterologous gene expression cassette containing the *bar* gene operably linked to the *Act1* promoter.

Construction of plasmid *rbcSP/rbcS* SP/*cbh1*//*pin3'*//*Act1 P/bar/nos3'*. The starting plasmid was pBR10-11 which contained the *crylA(b)* gene upstream of the *pin3'*. Between the *crylA(b)* and the *pin3'* is a DNA polylinker containing in the following order a *SmaI*, *BamHI*, *SpeI*, *XbaI*, *NotI*, and *EagI* restriction enzyme recognition site. The plasmid pBR10-11 (available from Silan Dai and Ray Wu, Department of Molecular Biology and Genetics, Biotechnology Building, Cornell University, Ithaca, New York 14853-2703) was digested with restriction enzymes *SpeI* and *XbaI* to produce a 9.2 kb DNA fragment. The 9.2 kb DNA fragment (pBR10-11/*SpeI/XbaI*/9.2 kb fragment) was purified by agarose gel electrophoresis.

The plasmid pB210-5a (available from William S. Adney, Mike Himmel, and Steve Thomas, National Renewable Energy Laboratory, 1670 Cole Boulevard, Golden Colorado 80401) containing the *cbh1* gene from *Trichoderma reesei* (*Trichoderma longibrachiatum*) was digested with *SpeI* and *XbaI*. The digested plasmid was electrophoresed on an agarose gel and a 1.8 kb fragment (pB210-5a/*SpeI/XbaI*/1.8 kb fragment containing *cbh1*) was

purified from the gel.

The above 9.2 kb and the 1.8 kb DNA fragments were ligated together using T4 DNA ligase to make plasmid "pBR10-11-cbh1" which was used to transform *E. coli* XL1 Blue. Transformed bacteria containing plasmid pBR10-11-cbh1 were identified by plating on LB agar gels containing ampicillin.

The plasmid pBR10-11-cbh1 was digested with *Sma*I and *Pst*I. The *Pst*I end was made blunt with mung bean exonuclease. The digested plasmid was electrophoresed on an agarose gel and the 2.8 kb DNA fragment containing cbh1 and pin3' was purified from the gel. The purified DNA fragment was designated "cbh1-pin3'/blunt-ended."

The plasmid pDM302 (Cao et al., Plant Cell Reports 11: 586-591 (1992)), shown in Figure 3, containing upstream of a *Cla*I site, a gene cassette consisting of the *bar* gene flanked by an upstream *Act*1 promoter and a downstream *nos*3', was digested with *Cla*I. The *Cla*I ends of the digested plasmid were made blunt with Taq DNA polymerase and the digested plasmid electrophoresed on an agarose gel. The digested plasmid was designated "pDM302/*Cla*I/blunt-ended."

The pDM302/*Cla*I/blunt-ended plasmid and the cbh1-pin3'/blunt-ended DNA fragment were ligated together using T4 DNA ligase to make plasmid "pDM302-cbh1-pin3'" which was used to transform *E. coli* XL1Blue. Transformed bacteria containing plasmid pDM302-cbh1-pin3' were identified by plating on LB agar gels containing ampicillin.

Plasmid pDM302-cbh1-pin3' was digested with *Spe*I, the ends made blunt with Taq DNA polymerase, and purified by agarose gel electrophoresis. The purified DNA fragment was designated "pDM302-cbh1-

pin3' / SpeI/blunt-ended."

Plasmid pRRI (available from Silan Dai and Ray Wu, Department of Molecular Biology and Genetics, Biotechnology Building, Cornell University, Ithaca, New York 14853-2703), which contains the rice *rbcS* small subunit gene, was digested with *PstI*. The *rbcS* promoter is flanked by *PstI* sites. The *PstI* ends were made blunt with mung bean nuclease and the 2 kb DNA fragment (rice *rbcS/PstI/blunt-ended*) containing the promoter was purified by agarose gel electrophoresis.

Rice *rbcSP/PstI/blunt-ended* and plasmid pDM-*cbh1-pin3' / SpeI/blunt-ended* were ligated using T4 DNA ligase to make *rbcSP/cbh1/pin3' // Act1P/bar/nos3'* which was then used to transform *E. coli* XL Blue. Transformed bacteria containing plasmid *rbcSP/cbh1/pin3' // Act1P/bar/nos3'* were identified by plating on LB agar gels containing ampicillin.

PCR was used to insert *NotI* sites into *rbcSP/cbh1/pin3' // Act1P/bar/nos3'*. These sites were used to insert the rice rubisco signal peptide in place of the *cbh1* signal peptide. The pRRI plasmid was the source of the rice rubisco signal peptide. It was also used as a PCR template to produce the PCR product containing the rice rubisco signal peptide flanked by *NotI* cohesive termini. The rice rubisco signal peptide and the *rbcSP/cbh1/pin3' // Act1P/bar/nos3'* plasmid were ligated together using T4 DNA ligase to make *rbcSP/rbcS SP/cbh1/pin3' // Act1P/bar/nos3'* which was then used to transform *E. coli* XL Blue. Transformed bacteria containing plasmid *rbcSP / rbcS SP/cbh1/pin3' // Act1P/bar/nos3'* were identified by plating on LB agar gels containing ampicillin.

Construction of plasmid *rbcSP/rbcS SP/e1/pin3' // Act1P/bar/nos3'*. Plasmid pMPT4-5

(available from William S. Adney, Mike Himmel, and Steve Thomas, national Renewable Energy laboratory, 1670 Colorado Boulevard, Golden, Colorado 80401) contains the el gene encoding endoglucanase I from *Acidothermus cellulolyticus* as a 3.7 kb *Pvu*I DNA fragment cloned into pGEM7 (Promega Corporation, Madison, Wisconsin). PCR was used to produce a DNA fragment containing the el gene flanked by *Asc*I recognition sites. Plasmid *rbcSP/cbh1/pin3'//Act1P/bar/nos3'* was also mutagenized by PCR to introduce *Asc*I sites flanking the *cbh1* gene. Next, the plasmid *rbcSP/cbh1/pin3'//Act1P/bar/nos3'* was digested with *Asc*I and the plasmid free of the *cbh1* gene was purified by agarose gel electrophoresis. The *Asc*I flanked el gene was ligated using T4 DNA ligase into the *rbcSP/cbh1/pin3'//Act1P/bar/nos3'* free of the *cbh1* gene to produce plasmid *rbcSP/el/pin3'//Act1P/bar/nos3'*, which then used to transform *E. coli* XL Blue. Transformed bacteria containing plasmid *rbcSP/el/pin3'//Act1P/bar/nos3'* were identified by plating on LB agar gels containing ampicillin.

PCR was used to insert *Not*I sites into *rbcSP/el/pin3'//Act1P/bar/nos3'*. These sites were used to insert the rice rubisco signal peptide in place of the *cbh1* signal peptide. The pRRI plasmid was the source of the rice rubisco signal peptide. It was also used as a PCR template to produce the PCR product containing the rice rubisco signal peptide flanked by *Not*I cohesive termini. The rice rubisco signal peptide and the *rbcSP/el/pin3'//Act1P/bar/nos3'* plasmid were ligated together using T4 DNA ligase to make *rbcSP/rbcSP/el/pin3'//Act1P/bar/nos3'* which was then used to transform *E. coli* XL Blue. Transformed bacteria containing plasmid *rbcSP/rbcSP/el/pin3'//Act1P/bar/nos3'* were identified by plating

on LB agar gels containing ampicillin.

Both heterologous gene expression cassettes are contiguous and the contiguous cassettes can be flanked by MAR sequences.

5

EXAMPLE 2

10 This example shows the construction of plasmids comprising a heterologous gene expression cassette comprising a DNA encoding a cellulase fusion protein. The plasmid constructs are shown in Table 2.

Table 2		
	Construct	Plasmid features
15	1 <i>rbcSP/cbh1/pin 3'</i>	<i>rbcSP</i> leaf-specific promoter driving cellulase cDNA of <i>T. reesei</i>
	2 <i>rbcSP/rbcS SP/cbh1/pin 3'</i>	The <i>rbcS</i> SP targets cellulase of <i>T. reesei</i> into maize chloroplasts
	3 <i>rbcSP/rbcS SP/syn-cbh1/pin 3'</i>	The <i>rbcS</i> SP targets modified cellulase of <i>T. reesei</i> into maize chloroplasts
	4 <i>CaMv35s/SSU/e1/nos3'</i>	The SSU targets the cellulase of <i>A. cellulolyticus</i> into maize chloroplasts
	5 <i>CaMv35s/VSP/e1/nos3'</i>	The VSP targets the cellulase of <i>A. cellulolyticus</i> into maize apoplasts
20	6 <i>CaMv35s/e1/nos3'</i>	No signal peptide

Abbreviations:

The term "syn-cbh1" refers to a *cbh1* gene that has been codon-modified for use in transformation of tobacco plants. It is available from .

25 The term "CaMV35s" refers to the cauliflower mosaic virus promoter.

The term "SSU" refers to the glycine max *rbcS* signal peptide. Glycine max is a soybean and not a rice variety.

The term "VSP" refers to the soybean vegetative storage protein beta signal peptide.

The remainder of the terms in Table 2 are the same as those for table 1.

5 Construct 1, which is shown in Figure 4, is plasmid pSMF13 which is plasmid pSK (Stratagene, La Jolla, California) which contains *cbh1* operably linked to the rice rubisco *rbcS* leaf-specific promoter which limits expression of the cellulase encoded by *cbh1* to
10 the cells of the leaves of the maize plant.

15 Construct 2, which is shown in Figure 5, is plasmid pSF15 which is plasmid pSK which contains *cbh1* operably linked to the rice rubisco *rbcS* leaf-specific promoter which limits expression of the cellulase encoded by *cbh1* to the cells of the leaves of the maize plant and a DNA encoding the *rbcS* SP which targets the cellulase to the chloroplasts.

20 Construct 3, which is shown in Figure 6, is like construct 2 except that the *cbh1* has been modified to decrease the GC content of the *cbh1* to an amount similar to the GC content of the tobacco plant genome. The nucleotide sequence of the modified *cbh1* (syn-*cbh1*) in plasmid pBI221 is set forth in SEQ ID NO:20.

25 Construct 4, which is shown in Figure 7, is plasmid pTZA8 which is plasmid pBI121 which contains the CaMV35s promoter, which is a constitutive promoter that is active in most plant tissues, to drive expression of *e1* which is operably linked to a DNA encoding the SSU signal peptide which targets the cellulase to the chloroplasts.

30 Construct 5, which is shown in Figure 8, is plasmid pZA9 which is similar to construct 4 except the signal peptide is encoded by DNA encoding the VSP signal peptide which targets the cellulase to the apoplasts.

Construct 6, which is shown in Figure 9, is plasmid pZA10 which is similar to construct 4 or 5 except that *el* is not operably linked to a DNA encoding a signal peptide that targets the cellulase to a plant organelle.

5

The constructs were prepared as follows.

10

First, the plasmid pRR1, which contains the rice *rbcS* gene was obtained from Ray Wu and Silan Dai, Cornell University. The rice rubisco (*rbcS*) small subunit was cleaved from pRR1 using *Eco*RI and *Eco*RV restriction sites to release a 2.1 kb DNA fragment containing the *rbcS*. The 2.1 kb DNA fragment was ligated into the plasmid pSK between the *Eco*R1 and *Eco*RV sites to produce plasmid pSMF8. The 2.1 kb DNA fragment provided the promoter for the *cbh1* constructs below.

15

Next, the *cbh1* gene was cloned downstream of *rbcS* promoter in plasmid pSMF8. First, a 1.7 kb DNA fragment containing the *cbh1* gene from *Trichoderma reesei* was isolated from plasmid pB210-5A (available from William S. Adney, Mike Himmel, and Steve Thomas, National Renewable Energy Laboratory; described in Shoemaker et al., Bio/Technology 1: 691-696 (1983)) by digesting with *Sal*I and *Xho*I. The ends of the 1.7 kb DNA fragment were made blunt end using DNA polymerase I (large fragment). The blunt-ended DNA fragment was cloned into plasmid pSMF8, which had been digested with *Bam*HI and the ends made blunt with DNA polymerase I, to make plasmid pSMF9.

20

Next, to complete the heterologous gene expression cassette, the *pin3'* transcription termination nucleotide sequence was inserted at the 3' end of the *cbh1* gene in plasmid pSMF9. The *pin3'* transcription termination nucleotide sequence was cleaved from pBR10-11 with *Pst*I. However, to remove the *pin3'* transcription termination nucleotide sequence from pBR10-11, a *Pst*I

RECEIVED
U.S. PATENT AND
TRADEMARK OFFICE
JULY 1988

site had to be introduced upstream of the *pin3'* transcription termination nucleotide sequence.

To generate the *PstI* site upstream of the *pin3'* transcription termination nucleotide sequence in pBR10-11, the pBR10-11 was digested with *NotI* and *XhoI* and a 70 bp multi-cloning site nucleotide sequence, which had been isolated from the pSK vector by digesting with *NotI* and *XhoI*, was cloned between the *NotI* and *XhoI* sites of the pBR10-11 to produce plasmid pSMF11. The *pin3'* transcription termination nucleotide sequence was then removed from pSMF11 by digesting with *PstI* to produce a 1 kb DNA fragment which was then cloned into the *PstI* site of pSK, which had been digested with *PstI*, to produce plasmid pSMF12. PSMF12 was then digested with *NotI* to produce a 1 kb DNA fragment containing the *pin3'* transcription termination nucleotide sequence. The 1 kb DNA fragment cloned into the *NotI* site downstream of the *cbh1* gene in pSMF9, which had been digested with *NotI*, to produce plasmid pSMF13 (construct 1 in Table 2).

Next, a DNA encoding a signal peptide which targets proteins to which it is attached to the chloroplasts was inserted upstream of the *cbh1* and in the same reading frame as the *cbh1*. Thus, a fusion protein is produced from translation of RNA transcribed from the *cbh1* DNA linked to the DNA encoding the signal peptide. DNA encoding the signal peptide (SP) was isolated from the *rbcS* in the pRR1 plasmid. Because there were no convenient restriction enzyme sites available which flanked the DNA encoding the SP for cloning, it was planned to PCR amplify that region containing the DNA encoding the SP using PCR primers with PCR primers that contained convenient restriction enzyme sites for cloning. At the end of the *rbcS*

5 promoter pSMF13 is a unique *AvrII* site and upstream of the first ATG of the *cbh1* gene is a unique *BsrGI*. A DNA encoding the SP that was flanked with an *AvrII* site on one end and a *BsrGI* site on the opposite end would be able to be cloned between the *AvrII* and *BsrGI* sites in pSMF13. That would place the DNA encoding the SP between the *rbcS* promoter and the *cbh1* gene and would enable a fusion protein containing the SP fused to the cellulase.

10 Therefore, PCR primers were synthesized using DNA sequences for the *AvrII* and *BsrGI* sites and the SP DNA sequences. The upstream PCR primer (SP1F) had the nucleotide sequence 5'-CCGCCTAGGCGCATGGCCCCCTCCGT-3' (SEQ ID NO:21) and the downstream PCR primer (SP3R) had the nucleotide sequence 5'-CGCTGTACACGCACCTGATCCTGCC-3' (SEQ ID NO:22). Plasmid pRR1 encoding the SP was PCR amplified with the PCR primers and the 145 bp amplified product was purified using 2% agarose gel. The purified 145 bp product was sequenced to confirm that the 145 bp amplified product contained the SP nucleotide sequences. The amplified product was digested with *AvrII* and *BsrGI* and cloned between the *AvrII* and *BsrGI* sites of pSMF13 digested with *AvrII* and *BsrGI* to produce plasmid pSMF14.

25 To produce pSMF15 which contains a *cbh1* gene codon-modified to decrease the GC content of the *cbh1* gene to an amount similar to the GC content of the tobacco genome, a synthetic *cbh1* (syn-*cbh1*) gene was obtained from plasmid pZD408 (available from Ziyu Dai, Pacific Northwest national Laboratory, 902 Battelle 30 Boulevard, Richland, Washington 99352). The syn-*cbh1* is a *cbh1* which had been codon-modified for use in tobacco plant transformations. The nucleotide sequence of syn-*cbh1* is set forth in SEQ ID NO:20. Plasmid pZD408 was linearized with *NcoI* and the ends made blunt. Then,

the blunt-ended pZD408 was digested with *Hind*III to remove the CaMV35S promoter. A 4.5 kb DNA fragment containing the syn-*cbh1* was isolated from the CaMV35S promoter by agarose gel electrophoresis. The 4.5 kb DNA fragment was dephosphorylated and the DNA fragment containing a blunt end and a *Hind*III end was named pZD408B.

Plasmid pSMF14 was digested with *Bsr*GI, the *Bsr*GI ends made blunt, and then pSMF14 was digested with *Hind*III to produce a DNA fragment containing the *rbcS* promoter with the DNA encoding the SP flanked by a blunt end and a *Hind*III end. The DNA fragment was purified by agarose gel electrophoresis and ligated to the pZD408B DNA fragment to produce plasmid pSMF15 (construct 3 of Table 2).

The heterologous gene expression cassettes are contiguous can be flanked by MAR sequences.

EXAMPLE 3

This example shows the construction of plasmids comprising a heterologous gene expression cassette comprising a DNA encoding a ligninase fusion protein and a heterologous gene expression cassette comprising a DNA encoding the *bar* gene. The constructs are shown in Table 3.

Table 3		
	Construct	Plasmid features
1	<i>rbcSP/ckg4/pin 3'//Act1 P/bar/nos 3'</i>	<i>rbcSP</i> leaf-specific promoter driving <i>ckg4</i> cDNA of <i>P. chrysosporium</i>
2	<i>rbcSP/ckg5/pin 3'//Act1 P/bar/nos 3'</i>	<i>rbcSP</i> leaf-specific promoter driving <i>ckg5</i> cDNA of <i>P. chrysosporium</i>
3	<i>rbcSP/rbcS SP/ckg4/pin 3'//Act1 P/bar/nos 3'</i>	The <i>rbcS</i> SP targets <i>ckg4</i> into maize chloroplasts

4	<i>rbcSP/rbcS SP/ckg5/pin 3'// Act1 P/bar/nos 3'</i>	The <i>rbcS</i> SP targets <i>ckg5</i> into maize chloroplasts
---	--	--

Abbreviations:

5 The terms "ckg4" and "ckg5" mean the ligninase cDNAs isolated from the basidiomycete *Phanerochaete chrysosporium*, SEQ ID NO:11 and SEQ ID NO:13, respectively. The codons for the 28 amino acid leader are deleted so that the expressed gene product remains inside the cells.

10 The remainder of the terms in Table 3 are the same as those for Table 1. All plasmid constructs contain the selectable marker gene (*bar*) driven by the rice actin 1 gene promoter. The rice actin gene and its promoter are disclosed in U.S. Patent 5,641,876 to McElroy et al.

15 Construct 1 contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg4* to the cells of the leaves of the maize plant.

20 Construct 2 contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg5* to the cells of the leaves of the maize plant.

25 Construct 3, which is shown in Figure 10, contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg4* to the cells of the leaves of the maize plant and further contains DNA encoding the *rbcS* SP which targets the ligninase to the chloroplasts.

30 Construct 4, which is shown in Figure 11, contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg5* to the cells of the leaves of the maize plant and

further contains DNA encoding the *rbcS* SP which targets the ligninase to the chloroplasts. All of the above constructs are adjacent to a heterologous gene expression cassette containing the *bar* gene operably linked to the *Act1* promoter. Both heterologous gene expression cassettes are contiguous and the contiguous cassettes can be flanked by MAR sequences.

EXAMPLE 4

This example shows the construction of plasmids comprising a heterologous gene expression cassette comprising a DNA encoding a ligninase fusion protein. The constructs are shown in Table 4.

15

Table 4		
	Construct	Plasmid features
1	<i>rbcSP/ckg4/pin 3'</i>	<i>rbcSP</i> leaf-specific promoter driving <i>ckg4</i> cDNA of <i>P. chrysosporium</i>
2	<i>rbcSP/ckg5/pin 3'</i>	<i>rbcSP</i> leaf-specific promoter driving <i>ckg5</i> cDNA of <i>P. chrysosporium</i>
3	<i>rbcSP/rbcS SP/ckg4/pin 3'</i>	The <i>rbcS</i> SP targets <i>ckg4</i> into maize chloroplasts
4	<i>rbcSP/rbcS SP/ckg5/pin 3'</i>	The <i>rbcS</i> SP targets <i>ckg5</i> into maize chloroplasts

20

The terms in table 4 are the same as those for Tables 1 and 3.

25 Construct 1, which is shown in Figure 12, is plasmid pSMF18 which is plasmid pSK which contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg4* to the cells of the leaves of the maize plant.

30 Construct 2, which is shown in Figure 13, is plasmid pSMF19 which is plasmid pSK which contains the

rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg5* to the cells of the leaves of the maize plant.

5 Construct 3, which is shown in Figure 14, is plasmid pMSF16 which is plasmid pSK which contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg4* to the cells of the leaves of the maize plant and further contains DNA encoding the *rbcS* SP which targets the ligninase to 10 the chloroplasts.

15 Construct 4, which is shown in Figure 15, is plasmid pSMF17 which is plasmid pSK which contains the rice rubisco *rbcS* leaf-specific promoter which limits expression of the ligninase encoded by *ckg5* to the cells of the leaves of the maize plant and further contains DNA encoding the *rbcS* SP which targets the ligninase to 20 the chloroplasts. The above heterologous gene expression cassettes can be flanked by MAR sequences.

25 The ligninase constructs shown in Table 4 are prepared as described below.

Two plasmids, pCLG4 and pCLG5, the former containing a cDNA clone encoding the ligninase gene *ckg4* and the latter containing a cDNA clone encoding the *ckg5* were obtained from Dr. C. Adinarayana Reddy, Department 30 of Microbiology and Public Health, Michigan State University and described in de Boer et al., Gene 6: 93-102 (1987), Corrigendum in Gene 69: 369 (1988). These ligninase cDNA clones were prepared from a white-rot filamentous fungus (*Phanerochaete chrysosporium*). The cDNAs for *ckg4* and *ckg5* had each been cloned into the *PstI* site of the pUC9 plasmid to make pCLG4 and pCLG5, respectively. The codons for the 28-amino acid leader sequence is deleted from both cDNAs before cloning so that expressed gene product remains inside the cell.

Plasmid pSMF16 is made as follows. The *ckg4* gene is removed from pCLG4 by digesting the plasmid with the restriction enzymes *Xba*I and *Bst*ΕΙΙ to produce a 1.2 kb DNA fragment containing the *ckg4* without the nucleotide sequence encoding the transit peptide. The *Bst*ΕΙΙ removes the nucleotide sequences encoding the transit peptide of the ligninase.

The ends of the DNA fragment containing the *ckg4* gene are made blunt and the blunt-ended DNA fragment is ligated into pSMF14 in which the *cbh1* has been removed by digesting with *Bsr*GI and *Xho*I and the ends made blunt to produce pSMF16.

Plasmid pSMF18 is made as follows. The nucleotide sequence encoding the *rbcS* signal peptide and *cbh1* are removed from pSMF14 by digesting pSMF14 with *Avr*II and *Xho*I instead of *Bsr*GI and *Xho*I. The ends of the digested pSMF14 are made blunt and the blunt-ended DNA fragment containing the *ckg4* gene, prepared as above, is ligated into the digested pSMF14 to make plasmid pSMF18.

Plasmid pSMF17 is made as follows. The *ckg5* gene is removed from pCLG5 by digesting the plasmid with the restriction enzymes *Xba*I and *Eag*I to produce a 1.2 kb DNA fragment containing the *ckg5* without the nucleotide sequence encoding the transit peptide. The *Eag*I removes the nucleotide sequences encoding the transit peptide of the ligninase.

The ends of the DNA fragment containing the *ckg5* are made blunt and the blunt-ended DNA fragment is ligated into pSMF14 in which the *cbh1* has been removed by digesting with *Bsr*GI and *Xho*I and the ends made blunt to produce pSMF17.

Plasmid pSMF19 is made as follows. The nucleotide sequence encoding the *rbcS* signal peptide and *cbh1* are

removed from pSMF14 by digesting pSMF14 with *AvrII* and *XhoI* instead of *BsrGI* and *XhoI*. The ends of the digested pSMF14 are made blunt and the blunt-ended DNA fragment containing the *ckg5* gene, prepared as above, is
5 ligated into the digested pSMF14 to make plasmid pSMF19.

EXAMPLE 5

This example shows the transformation of maize
10 multi-meristem primordia via Biolistic bombardment with
the plasmid constructs of Examples 1-4, regeneration of
the transgenic plants, confirmation of the integration
of the plasmid constructs into the plant genome, and
confirmation of the expression of the cellulase or
15 ligninase fusion proteins in the transgenic plant. For
transformations with the constructs of Examples 2 and 4,
which do not contain a selectable marker, a selectable
marker comprising the *bar* gene in the plasmid pDM302
(Cao et al., Plant Cell Reports 11: 586-591 (1992)) is
20 cotransfected into the cells with the plasmid containing
the ligninase or cellulase heterologous gene expression
cassette.

Maize seeds have been germinated in Murashige
and Skoog (MS) medium (Murashige and Skoog, Physiol.
Plant 15: 473-497 (1962)) supplemented with the
25 appropriate growth regulators (Zhong et al., Planta 187:
490-497 (1992)). Shoot meristems have been dissected
and cultured for 2-3 weeks until an initial
multiplication of meristem have been produced for
bombardment.

30 The multi-meristem primordia explants are
bombarded with tungsten particles coated with particular
plasmids of Example 1 or 3 or with particular plasmids
of Example 2 or 4 along with the plasmid containing the
heterogenous gene expression cassette containing the *bar*
35 gene. The bombarded explants are gently transferred

onto meristem multiplication medium for further multiplication, about 6 to 8 more weeks. This step is required to reduce the degree of chimerism in transformed shoots prior to their chemical selection.

5 Shoots are transferred to the above medium containing 5 to 10 mg per liter glufosinate ammonium (PPT) selectable chemical for another 6 to 8 weeks. Chemically selected shoots are rooted in rooting medium containing the same concentration of PPT. Plantlets are transferred to

10 pots, acclimated, and then transferred to a greenhouse.

When the plantlets or shoots are small, the quantity of transgenic plant material is insufficient for providing enough DNA for Southern blot hybridization; therefore, polymerase chain reaction (PCR) is used to confirm the presence of the plasmid constructs the plantlets. The amplified DNA produced by PCR is resolved by agarose or acrylamide gel electrophoresis, transferred to membranes according standard Southern transfer methods, and probed with the appropriate DNA construct or portion thereof according to standard Southern hybridization methods. Those shoots or plantlets which show they contain the construct in its proper form are considered to have been transformed. The transformed shoots or plantlets are grown in the greenhouse to produce sufficient plant material to confirm that the plasmid constructs has been properly integrated into the plant genome. To confirm proper integration of the plasmid constructs into the plant genome, genomic DNA is isolated from the greenhouse grown transgenic plants and untransformed controls and analyzed by standard Southern blotting methods as in Zhong et al., Plant Physiology 110: 1097-1107 (1996); Zhang et al., Theor. Appl. Genet. 92: 752-761 (1996); Zhang et al., Plant Science 116: 73-84 (1996); and, Jenes et al., In *Transgenic Plants*. Vol. 1.

Kung, S-D and Wu, R (eds.). Academic Press, San Diego, CA. pp. 125-146 (1992).

To confirm expression of the ligninase or cellulase fusion protein, total cellular RNA is isolated
5 from the greenhouse grown plant tissues as described in
Zhong et al., Plant Physiology 110: 1097-1107 (1996).
The mRNA encoding the cellulase or ligninase fusion
protein is detected by RNA Northern blot analysis using
the same probes used for the Southern blot analyses.
10 Briefly, the RNA is electrophoresed on a denaturing
formaldehyde agarose gel, transferred to nitrocellulose
or nylon membranes, hybridized to the appropriate
ligninase or cellulase probe, and then exposed to X-ray
autoradiology film. The hybridization bands are scanned
15 using a densitometer which enables determination of the
expression level of the specific mRNA.

Translation of the mRNA is confirmed by
Western blot analysis according to the standard methods
of Towbin et al., Proc. Natl. Acad. Sci. USA 76: 4350
20 (1979) and Burnette, Anal. Biochem. 112: 195 (1981)
using antibodies specific for ligninase or cellulase.

EXAMPLE 6

Transgenic maize containing both a ligninase
25 and a cellulase fusion protein is made by crossing-
breeding the abovementioned transgenic plants one of
which contains *cbh1* or *el* stably integrated into the
genome and the other of which contains *ckg4* or *ckg5*
30 stably integrated into the genome using the method
provided in (Zhong et al, Theor. Appl. Genet. 92: 752-
761, (1996); Zhong et al, Plant Physiol. 110: 1097-1107,
(1996); Zhong et al, Planta, 187: 483-489, (1992)).
Transgenic plants that carry a low copy number of the
35 DNA encoding the ligninase or cellulase fusion proteins
are used for cross-breeding.

Briefly, transgenic maize plants that produce the ligninase fusion protein are made as disclosed in Example 5 to make a first transgenic plant and transgenic maize plants that produce the cellulase fusion protein are made as disclosed in Example 5 to make a second transgenic plant. The first and second transgenic plants are cross-pollinated to create a transgenic plant which produces both a ligninase and a cellulase fusion protein. The progeny are analyzed for homozygosity and transgenic plants that are homozygous for both the ligninase gene cassette and the cellulase gene cassette are selected for further propagation for seeds.

The progeny in the above crosses are used in subsequent crosses to produce transgenic maize with both ligninase gene cassettes and one, two, or three cellulase gene cassettes or transgenic maize with two or three cellulase gene cassettes and one ligninase gene cassette.

EXAMPLE 7

Production levels and activity of the cellulase fusion protein in transgenic maize made as in Example 5 or 6 is determined as follows.

Cellulase activity in transgenic maize is first assayed by standard methods (Ghose. In *Analytical Method B-304, rev. A, IUPAC Commission on Biotechnology. A short Report (1984)*) based on the time course assay for hydrolysis of a pre-weighed sample of filter paper at pH 4.8-5.2 and temperature of 50° C. While the filter paper assay is a standard substrate for cellulase activity, results using the filter paper assay are not particularly representative of the actual activity of the cellulase in plant materials containing cellulose, hemicellulose, and other sugars or sugar polymers.

Therefore, a more accurate method for determining cellulase activity is used.

Plant material is ground and the ground material is suspended to a concentration of up to about 5% in 0.05 M citrate buffer at pH 4.8 and incubated with shaking at 50° C. Over a 48 hour time period, samples are removed at intervals of 0, 1, 3, 12, 24, and 48 hours. A minimal amount of sodium azide, about 0.05%, is added to the citrate buffer during incubation to 10 control microbial activity. For analysis by high pressure liquid chromatography (HPLC), the supernatant fraction of each sample is removed, capped, and heated to inactivate the enzymes. The inactivated supernatant fraction is filtered through a syringe filter and 15 analyzed by HPLC to measure the glucose, cellobiose, and xylose content of the samples according to established methods (Dale et al., Biosource Technol. 56: 11-116 (1996)).

Cellulase activity is manifested by an 20 increasing level of glucose, xylose and/or cellobiose levels in the supernatant fractions during the 48 hour period. The control for the above assay is to treat samples from non-transgenic plants with varying amounts 25 of commercially available cellulase enzymes such as CYTOLASE 300 which is a cellulase from Genencor, Inc. and NOVOZYME 188 which is a cellobiose from Novo Laboratories, Inc. to confirm that the ground plant material is susceptible to hydrolysis.

30

EXAMPLE 8

Comparison of cellulase activity in transgenic maize prepared as in Example 5 or 6 treated to enhance cellulose accessibility.

Generally, cellulose and hemicellulose in 35 plant material are not very accessible to hydrolytic

enzymes such as cellulase. Therefore, it is possible that even if the cellulase fusion protein is produced in the transgenic plants of the present invention, its cellulase activity would not be measurable. Therefore, 5 to demonstrate accessibility, samples of the transgenic maize plants of the present invention are treated by the ammonia fiber explosion process to increase cellulose and hemicellulose accessibility (Dale et al., Biosource technol. 56: 11-116 (1996)). Samples treated are 10 analyzed as in Example 3.

In previous experiments with coastal Bermuda grass, the ammonia fiber explosion process disrupted the plant cell walls sufficiently to permit over 80% extraction of plant protein, compared with less than 30% extraction under the same conditions prior to ammonia treatment (de la Rosa et al., Appl. Biochem. Biotechnol. 45/46: 483-497 (1994)). The process increased the hydrolytic effectiveness of the added cellulases by at least six-fold (Dale et al., Biosource Technol. 56: 11-116 (1996)). Thus, it is expected that the ammonia fiber explosion process helps release cellulase from the transgenic maize chloroplasts and will also increase the access of the cellulase released to the cellulose in the plant material.

25

EXAMPLE 9

Production levels and activity of the ligninase fusion protein in transgenic maize made as in Example 5 or 6 can be determined as follows.

30 Maize leaves from the transgenic maize made as in Examples 5 or 6 are ground using a pestle and mortar. Chloroplasts are isolated from leaves of transgenic plants by Ficoll (Pharmacia) gradient centrifugation and ground as above.

35 The ground materials (leaves, grains,

chloroplasts) are suspended in 50 mM L-tartrate buffer (pH 4.5), mixed well by vortexing, and centrifuged for 10 minutes at 14,000 rpm (16,000 x g) at 4° C and the supernatant fraction tested for lignin peroxidase (LIP) 5 activity as described in Tien et al., *Meth. Enzymol.* 161: 238-249 (1988). The LIP assay measures the production of veratraldehyde (as an increase in absorbance at 310 nm) from veratryl alcohol (substrate) in the presence of hydrogen peroxide. Control assays 10 are done on non-transgenic maize seeds to measure endogenous peroxidase activity. The assay is sensitive and is able to detect very low levels of lignin peroxidase activity, e.g., conversion of 0.1 mmole substrate per minute per liter of test sample.

15 Soluble protein content is determined by the Bradford procedure (Bradford, *Anal. Biochem.* 72: 248-254 (1976)) using bovine serum albumen (BSA) as the standard. LIP enzyme in the extracted fluid is purified by Fast Protein liquid Chromatography (FPLC) analysis 20 using the Mono Q anion exchange system (Pharmacia) and a gradient of 0 to 1 M Na-acetate to elute the various isozymes (Yadav et al., *Appl. Environ. Microbiol.* 61: 2560-2565 (1995); Reddy et al., *FEMS Microbiol. Rev.* 13: 137-152 (1994)). The relative activity, yield, pH 25 optimum, stability, and other characteristics of the LIP in the transgenic plant are compared to that determined for the LIP isolated from the fungus. Furthermore, ground maize seeds or leaf extracts containing the LIP 30 is used to treat various lignocellulosic feeds in small laboratory reactor systems and the extent of delignification can be analyzed per established procedures (Van Soest et al., *Assoc. Off. Anal. Chem. J.* 51: 780-785 (1968)).

35 Detection of ligninase mRNA is by isolating the mRNA from the transgenic plants as above, resolving

the mRNA by denaturing RNA gel electrophoresis, transferring the resolved mRNA to membranes, and probing the membranes with *ckg4* or *ckg5* cDNA probes.

5 Western blots are performed to determine whether the LIP protein is in an active or inactive form. The total protein from the transgenic plants is resolved by SDS-polyacrylamide gel electrophoresis and transferred to membranes. The membranes are probed with antibodies to LIP H2 (*ckg4*) or LIP H10 (*ckg5*).

10

15

While the present invention is described herein with reference to illustrated embodiments, it should be understood that the invention is not limited hereto. Those having ordinary skill in the art and access to the teachings herein will recognize additional modifications and embodiments within the scope thereof. Therefore, the present invention is limited only by the Claims attached herein.

SEARCHED
INDEXED
SERIALIZED
FILED